

Review

Maxime Bodak, Jian Yu and Constance Ciaudo*

Regulation of LINE-1 in mammals

Abstract: Transposable elements (TEs) are mobile DNA elements that represent almost half of the human genome. Transposition of TEs has been implicated as a source of genome evolution and acquisition of new traits but also as an origin of diseases. The activity of these elements is therefore tightly regulated during the life cycle of each individual, and many recent discoveries involved the genetic and epigenetic mechanisms in their control. In this review, we present recent findings in this field of research, focusing on the case of one specific family of TEs: the long-interspersed nuclear elements-1 (LINE-1 or L1). LINE-1 elements are the most representative class of retrotransposons in mammalian genomes. We illustrate how these elements are conserved between mice and humans, and how they are regulated during the life cycle. Additionally, recent advances in genome-wide sequencing approaches allow us not only to better understand the regulation of LINE-1 but also highlight new issues specifically at the bioinformatics level. Therefore, we discuss the state of the art in analyzing such bioinformatics datasets to identify epigenetic regulators of repeated elements in the human genomes.

Keywords: bioinformatic challenges; LINE-1 regulation; repeat elements; transposable elements.

DOI 10.1515/bmc-2014-0018

Received June 23, 2014; accepted August 19, 2014

List of abbreviations: 5caC, 5-carboxylcytosine; 5fC, 5-formylcytosine; 5hmC, 5-hydroxymethylcytosine; 5mC, 5-methylcytosine; AID, activation-induced deaminase; APOBEC, apolipoprotein B mRNA-editing enzyme complex; ATPase, adenosine triphosphatase; ChIP-seq, chromatin immunoprecipitation followed by sequencing; DNA, deoxyribonucleic acid; Dnmt, DNA methyltransferase; dpc, day postcoitum; ERV, endogenous retrovirus; FOA, fetal oocyte attrition; GSE, gonad-specific expression; HELP-seq, HpaII-tiny fragment enrichment by ligation-mediated PCR coupled to massively parallel DNA sequencing; HIV, human immunodeficiency virus; hnRNP, heterogeneous nuclear ribonucleoprotein; LINE-1 or L1, long-interspersed nuclear elements-1 or long interspersed element; Lsh, lymphoid-specific helicase; LTR, long terminal repeat; MbD, methyl-CpG-binding domain; MHR, multiple-hit read; miRNA, microRNA; mRNA, messenger RNA; NHP, non-human primate; NPC, neuronal progenitor cell; nt, nucleotide; ORF, open reading frame; PABP, poly(A) binding protein; PGC, primordial germ cell; piRNA, PIWI-interacting RNA; RDC, Rhino, Deadlock, and Cutoff; RISC, RNA-induced silencing complex; RNA, ribonucleic acid; RNAi, RNA interference; RNP, ribonucleoprotein; RSEM, RNA-Seq by expectation maximization; siRNA, small interfering RNA; shRNA, short hairpin RNA; TDRD, Tudor domain containing; TEs, transposable elements; TET, ten-eleven translocation; TF, transcription factor; TPRT, target primed reverse transcription; UHR, unique-hit read; UTR, untranslated region.

***Corresponding author: Constance Ciaudo**, Swiss Federal Institute of Technology Zurich, Department of Biology, Institute of Molecular Health Sciences, HPL G32.1, Otto-Stern-Weg 7, CH-8093 Zurich, Switzerland, e-mail: cciaudo@ethz.ch

Maxime Bodak: Swiss Federal Institute of Technology Zurich, Department of Biology, Institute of Molecular Health Sciences, Life Science Zurich Graduate School, Molecular Life Science Program, HPL G28, Otto-Stern-Weg 7, CH-8093 Zurich, Switzerland

Jian Yu: Swiss Federal Institute of Technology Zurich, Department of Biology, Institute of Molecular Health Sciences, Life Science Zurich Graduate School, Molecular and Translational Biomedicine Program, HPL G28, Otto-Stern-Weg 7, CH-8093 Zurich, Switzerland

Introduction

Since their discovery by Barbara McClintock in the 1940s (1), transposable elements (TEs) still continue to enthuse and inspire researchers. Significant progress has been made in understanding the biology of TEs, which is reflected by the increasing number of published studies. The popularity of these classes of DNA stems from their ability to translocate to a new location in a genome, explaining why TEs are also called ‘jumping genes’ or ‘mobile elements’. Moreover, TEs account for 46% of the human genome (2) and about 39% of the mouse genome

(3). They are divided into two classes, according to a classification proposed by Finnegan in 1989 (4) and revisited by Wicker et al. in 2007 (5), based on their mechanism of transposition: class I retrotransposons require an RNA intermediate involving a ‘copy-and-paste’ mechanism (6) and class II DNA transposons move by a ‘cut-and-paste’ mechanism (7). Retrotransposons represent about 42% of the human and 38% of the mouse genome, whereas DNA transposons account for only 3% and 1%, respectively (2, 3). In this review, we will focus on retrotransposons and more specifically on the regulation of long-interspersed nuclear elements-1 (LINE-1 or L1), a subclass of this family. L1 is the most representative class of retrotransposons in the mammalian genome, representing 17% of the human (2) and almost 19% of the mouse genome (3). They belong to the autonomous non-LTR retrotransposon category, owing to their ability to encode the proteins required for their mobility. L1 elements are about 6–7 kb in length and contain a 5′ untranslated region (5′UTR) with an internal promoter activity, two open reading frames (ORF1 and ORF2), and a 3′UTR that ends with a poly(A) tail (Figure 1A). ORF1 encodes a 40 kDa protein with RNA binding activity (8), whereas ORF2 encodes a 150 kDa

protein with a reverse transcriptase (9) and an endonuclease domain (10) (Figure 1A). Active L1 elements must be able to perform a complete retrotransposition cycle that includes the following steps in chronological order: transcription of L1 RNA, export into the cytoplasm, translation of ORF1 and ORF2, association of L1 RNA with ORF1 and ORF2 proteins to form ribonucleoprotein (RNP) particles (11), return to the nucleus, reverse transcription, and integration at a new genomic location – a one-step process also called target primed reverse transcription (TPRT) (12) (Figure 1B). During this retrotransposition event, the 5′UTR is frequently truncated (2) and the L1 sequence is prone to inversion (13). These are the two main reasons for the accumulation of inactive L1 copies in the genome. Indeed, the number of L1 fragments in the human genome has been estimated at 516 000 elements (2); however, only 80–100 of them have kept their mobility capability (14). Similarly, the mouse genome contained almost 600 000 L1 fragments (3); however, previous studies showed that only 3000 full-length elements are potentially able to retrotranspose (15, 16). These active L1 copies can deeply influence the genome in numerous ways, beneficial and detrimental (17). L1 retrotransposition can act as a

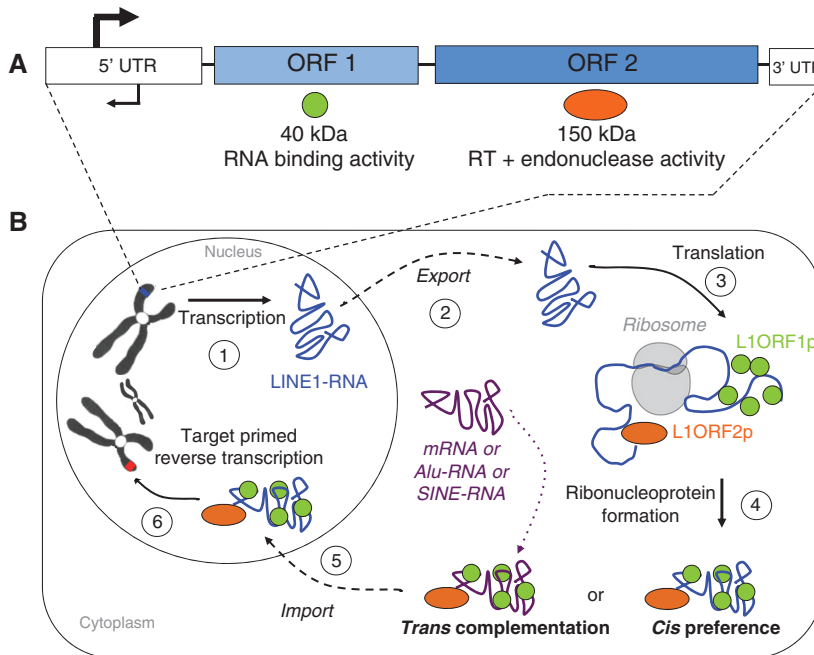


Figure 1 Schematic representation of L1 structure and retrotransposition cycle.

(A) L1 anatomy – an L1 is composed of a 5′UTR, two ORFs each coding for a protein [ORF1p (40 kDa) and ORF2p (150 kDa)], and a 3′UTR. (B) L1 retrotransposition cycle – during a complete retrotransposition cycle, an L1 element is transcribed (1) and the resulting L1 RNA is exported out of the nucleus into the cytoplasm (2). Then, the bicistronic RNA is translated (3) into ORF1 and ORF2 proteins, which bind the L1-RNA to form RNP particles (4): this is the *cis* preference. ORF1p and ORF2p can also bind other RNAs (such as cellular mRNA, Alu-RNA, or SINE-RNA): this is called the *trans* complementation phenomenon. Finally, RNP complexes are imported into the nucleus (5) where the RNA can be integrated at a new genomic location by the process of TPRT (6).

mutagen by inserting into exons, inducing aberrant splicing or exon skipping by inserting into introns, and less commonly, they can give rise to a gene-breaking phenomenon (18). This last example could generate distinct transcription units (19), which can lead to sporadic cases of diseases (18). Thereby, since their discovery, almost 65 diseases-causing mutations in humans have been attributed to L1-mediated retrotransposition. Moreover, L1 elements are able to mediate the retrotransposition of other RNAs, a phenomenon called *trans* complementation (Figure 1B). In this case, L1 can mobilize non-autonomous retrotransposons such as short interspersed elements (SINEs) and Alu elements (20) or process pseudogenes from cellular mRNAs (21), and therefore participate in genome expansion (18). Here, however, we will focus on the regulation of L1 transcription and retrotransposition, as the impact of new L1 insertions on the genome has already been extensively reviewed (12, 17, 18). In humans, the current rate of L1 retrotransposition has been estimated between 1 insertion out of 20 and 1 out of 200 live births depending on the analysis method (18). L1 elements have been described to be expressed in the germ line, during early development and in some somatic tissues, as well as in tumor cells (22). In this review, we will focus on the regulatory mechanisms involved in the control of these TEs during the life cycle in mammals, from gamete precursors to somatic cells through the blastocyst stage. Our goal is to provide an overview of what is known about L1 regulation itself, the latest discoveries in the field, and the tools available to analyze such repeat elements. First, we will present a comparison of active L1 5'UTR from humans and mice, the two main mammalian systems used to study L1 regulation in mammals. Subsequently, we will describe the numerous regulatory mechanisms of L1 that are presently known and the latest hypothesis. Finally, we will present the bioinformatics tools used to study repeated elements as well as the problems related to the high number of copies present in the genomes. We will conclude our review with potential future research directions.

Comparison of human and mouse active LINE-1 elements

Numerous studies have been conducted using mouse L1 constructions in human cell lines and *vice versa*, and have shown that human L1 are able to retrotranspose in the mouse, *in vitro* and *in vivo* (23, 24). At the same time, other studies employ artificial L1 constructions to better understand the mechanisms that can rule their regulation (25,

26). However, even if mouse and human L1 share common features concerning their biology, they also differ on many points, putting into question the conservation of the L1 regulatory mechanisms between the two mammalian systems used to study this class of TEs.

Differences in mobilization activity

First, mouse and human genomes differ considerably concerning the number of L1 active copies (27): the percentage of active elements according to the number of copies in the genome is 25 times higher in mice (0.5%) than in humans (0.015–0.020%) (2, 3, 15, 16). This observation suggests a decline of L1 activity in humans. Comparative gene expression analysis performed on human and non-human primate (NHP) induced pluripotent stem cells revealed regulation differences of L1 with an increase of their mobility in the NHP genomes (28). Regulation differences between human and NHPs support the idea that mouse and human L1 regulatory mechanisms can diverge as well.

Differences in structure

Furthermore, mouse and human L1 elements also differ in their structure: mouse L1 are approximately 7 kb in length (17), whereas human L1 are around 6 kb (12). Moreover, the two ORFs of human L1 are in frame and separated by a 67–70 bp non-coding spacer region containing a stop codon in the three reading frames. In mouse, the two ORFs are also separated by a spacer region (29–115 bp) (29) and they have different reading frames. Finally, unlike human L1, the 5'UTR of mouse L1 contains variable numbers of monomers, which are tandem repeat units of 205–210 bp with an intrinsic promoter activity (30). This allows the definition of several L1 subfamilies in the Repbase database (31), among them three are proved to be active in mice: A (32), Tf (15), and Gf (16).

Focus on the 5'UTR, a key region for L1 regulation

L1 5'UTR contains an internal promoter activity independent of upstream sequences (33), as well as an antisense promoter (34–36), making the 5'UTR a key region in the regulation of these elements and in the definition of an active element. Moreover, human and mouse L1 5'UTR have different putative transcription factor (TF) binding sites (37). Considering these evidences, we chose to focus

on the 5'UTR of active elements to gain deeper insight into the comparison between human and mouse L1. As we previously mentioned in the Introduction, the human genome contains around 80–100 active L1, called retrotransposition competent L1 (RC_L1) (14). In the literature, RC_L1 are also denoted as L1Hs or L1-PA1 (31). It has been recently shown that the 5'UTR of RC_L1 could adopt secondary structures that can be recognized and processed by protein complexes (38). Here, we present a comparison of the sequences and potential secondary structures of 5'UTR from active L1 subfamilies, including A, Tf, and Gf in mouse and RC_L1 (Figure 2). To investigate the difference among these four regions, we considered the consensus sequence of each 5' UTR element (31) and used two different approaches: (i) drawing the phylogenetic tree using a neighbor-joining algorithm to reveal the evolutionary relation (39) (Figure 2A), and (ii) performing sequences and structure alignments by using LocARNA (40) (Figure 2B). As expected, mouse L1 5'UTR from the different subfamilies clustered into the same group with both methods.

According to these trees, L1-Tf and L1-Gf 5'UTR appear to be the most similar regions, whereas RC_L1 5'UTR is, in both cases, the most divergent (Figure 2A,B). Then, we compared the predicted secondary structures of these four L1 5'UTR by RNAfold (41), and we found a substantial divergence not only between the human and mouse L1 5'UTR region but also among the mouse L1 5'UTR subfamilies (Figure 2C). These differences suggest that the mechanisms involved in the regulation of active mouse and human L1 can be potentially distinct, even between subfamilies.

L1 regulation during the mammalian life cycle

It has been estimated that in humans, one new genomic L1 insertion appears in every 10–250 individuals and that the proportion of L1 retrotransposition causing disease could be around 0.07% (12). In mice, the number of active L1 has

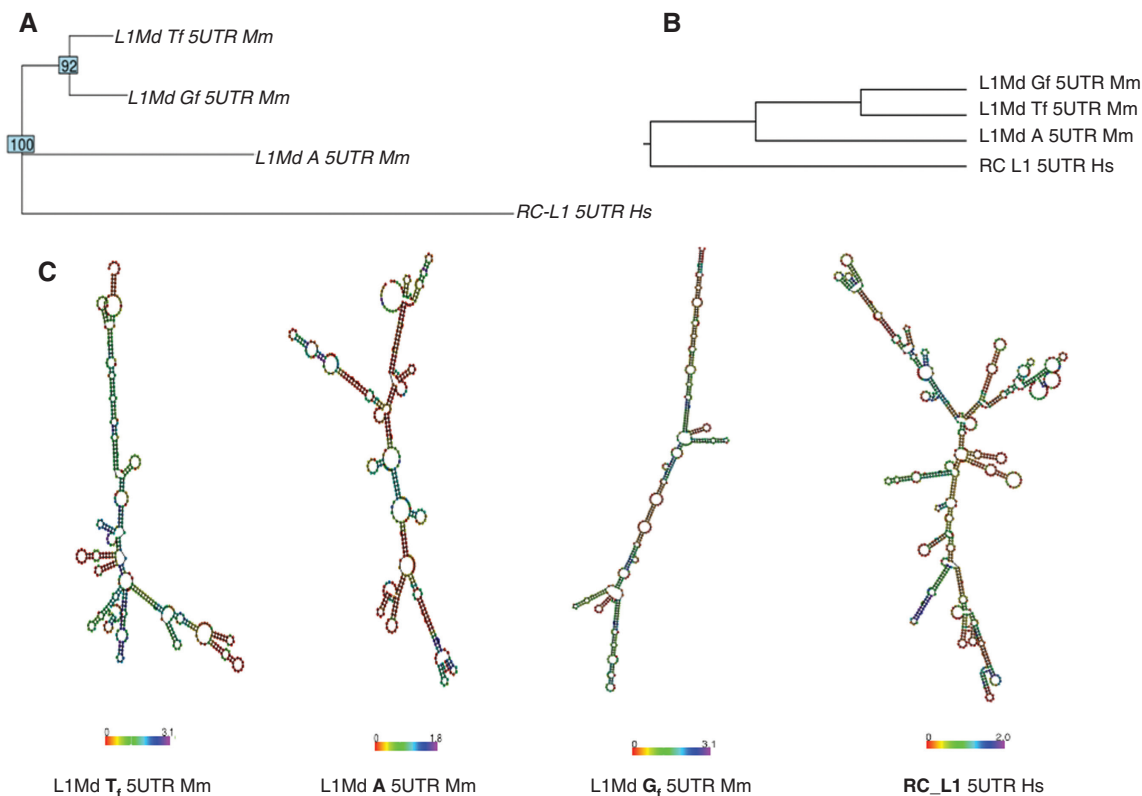


Figure 2 Sequences and structures comparison of human (RC_L1) and mouse (Gf, Tf, A) L1 5'UTRs.

(A) Phylogenetic tree generated with neighbor-joining algorithms using the Kimura model. The numbers on the nodes indicate the percentage of the labeled node present in 93 bootstrap replicates. (B) Multiple sequence and secondary structure alignment by LocARNA. The height of the tree corresponds to the LocARNA scores, where lower scores/heights indicate a closer distance. (C) Secondary structures were predicted by RNAfold. The color represents the entropy for each base. L1Md Tf 5'UTR Mm, Tf mouse L1 subfamily; L1Md A 5'UTR Mm, a mouse L1 subfamily; L1Md Gf 5'UTR Mm, Gf mouse L1 subfamily; RC_L1 5'UTR Hs, human retrotransposition competent L1.

been estimated using cell culture assays to a total of 3100. In terms of ratio of active copies per total copies present in the genome, the proportion of operational L1 in mice is 25 times higher than in humans. Owing to their high mutagenic potential, these elements must be strongly regulated during the development and adult life of each individual. In this part of the review, we present the state of the art of the L1 regulatory mechanisms described in the literature. As a guideline, this section is divided according to the different steps of the mammalian life cycle; thus, we first describe the pathways regulating L1 in the germ line, after fertilization and to the blastocyst stage. Subsequently, we depict which mechanisms are involved in sustaining the control in somatic cells, and finally, we discuss about the particular cases of L1 reactivation in somatic cells.

Epigenetic regulation through DNA methylation is one of the main mechanisms, which has evolved to play a role in the defense against TEs through transcriptional silencing

(42). During the mouse life cycle, L1 promoters undergo two waves of partial demethylation: one occurs during the specification and migration of primordial germ cells (PGCs) and the other during the preimplantation stage (43) (Figure 3A). However, DNA methylation is not the only L1 repressor system used to control their transcription and retrotransposition. Therefore, it has been proposed that overlapping epigenetic mechanisms evolved to control the expression of TEs in eukaryotic cells (44). Such example of multiple regulations of L1 subtypes has been recently demonstrated in human and mouse embryonic stem cells (mESCs) (45) (see ‘Regulation in the blastocyst’).

Regulation in the germ cell lineage

The majority of TEs spread in a population according to a vertical mode of inheritance (46). This transmission mode

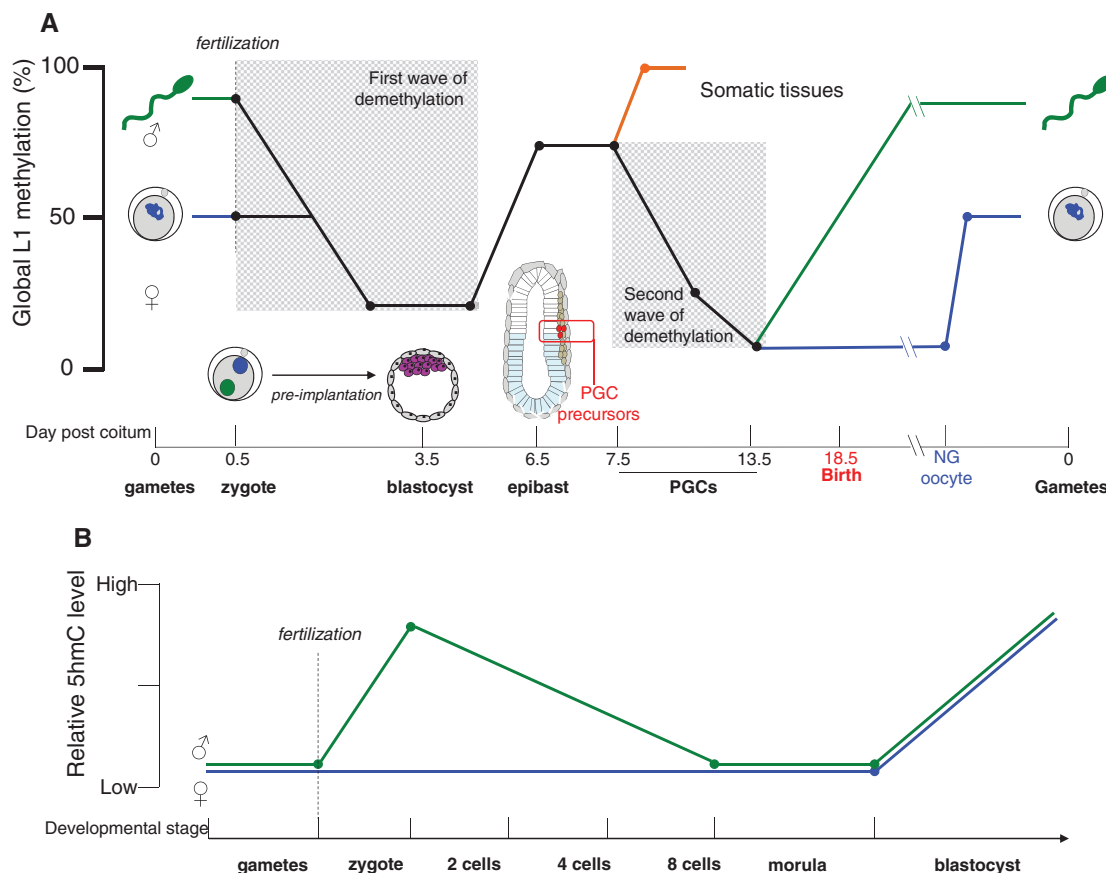


Figure 3 Global L1 methylation and 5hmC level during several steps of mammalian development and associated L1 regulatory mechanisms. (A) Schematic representation of the global L1 methylation level of the mouse genome. During embryonic development, mouse L1 undergoes two main waves of demethylation. The first occurs after fertilization: from the zygote stage (0.5 dpc) until the blastocyst stage (3.5 dpc). The second occurs during the migration of the PGC precursors from the proximal epiblast to the gonad precursors (from 7.5 to 13.5 dpc). PGC, primordial germ cell; NG oocyte, non-growing oocyte. Adapted from ref. (43). (B) Schematic representation of 5-hydroxymethylcytosine dynamics in paternal and maternal genomes during preimplantation development. Adapted from ref. (47).

requires an activity in gamete precursors, making the germ line a pertinent developmental context to study L1 regulation. In mice, at 6.5 days postcoitum (dpc), cells from the proximal epiblast will differentiate into PGCs and migrate to reach the precursor of the gonads. These cells will form the germ line. During this process, the early developing germ line acquires pluripotency, an event coinciding with a relaxation of the epigenetic repression of the genome (43) (Figure 3A). This developmental window is indeed an opportunity for L1 invasion of the host genome.

Role of DNA methylation

DNA methylation is known to be involved in various biological processes as a regulator of gene expression, X-chromosome inactivation, genomic imprinting, and also in the silencing of TEs (47, 48). The reestablishment of the methylation status of L1 elements requires DNA methyltransferase proteins involved in the DNA methylation machinery. It has been shown that mutant mouse embryos for the maintenance DNA-methyltransferase DNMT1 lose the methylation of several types of TEs (49). A similar phenotype has been observed for the double-mutant mouse for *Dnmt3a* and *Dnmt3b* genes encoding for two *de novo* DNA methyltransferases (50).

Studies from Bestor laboratory highlighted yet another player in this process: DNMT3L, the DNMT3A cofactor that is strictly required for L1 methylation in the male germ line (51). DNMT3L stabilizes the active conformation of DNMT3A, thereby increasing the efficiency of methyl group transfer onto target site and facilitating *de novo* methylation (52). Problems in completing this procedure involve the high expression level of L1 in germ cells, and also sterility, similar to the phenotype observed in *Dnmt3A* and *Dnmt3L* male mutant mice (51). The *Dnmt3L* gene is also evolutionarily correlated to the germ line protection against TEs, as it emerged in eutherian mammals around 150 million years ago coinciding with an important TE expansion in mammalian genomes (53).

Furthermore, proteins assisting the DNA methylation process also play a role in TE regulation. For example, the protein UHRF1, for ubiquitin-like containing PHD and RING finger domains 1, is required by DNMT1 to load onto hemimethylated DNA strands. *Uhrf1* inactivation in mice results in decreased methylation of L1 (54). The protein lymphoid-specific helicase (LSH) should also be mentioned in this context. This member of the SNF2 family of chromatin remodeling ATPases supports the access of DNMTs to the DNA. Initially, LSH was described to be directly recruited to L1 elements in mice (55). More

recently, Meehan laboratory developed a new technology called HELP-seq (for HpaII tiny fragment enrichment by ligation-mediated PCR coupled to massively parallel DNA sequencing) to profile the methylation of TEs in *Lsh* and *Dnmt3b* mutant fibroblasts (56). They observed hypomethylation of L1 in *Lsh*^{-/-} and *Dnmt3b*^{-/-} cells, which is consistent with previous investigations implicating LSH as a recruitment protein for *de novo* methylation. Nevertheless, they also observed differences of methylation between L1 subfamilies and a greater hypomethylation in the absence of LSH than DNMT3B, suggesting DNMT3B-independent roles for LSH in repeat methylation.

Finally, a recent study showed that in addition to DNA methylation, the PIWI-interacting RNA (piRNA) pathway is required to maintain a high level of repressive H3K9me3 on L1 elements in mouse germ cells (57).

Role of the piRNA pathway

Studies performed on germ cells from *Drosophila* depicted a new TE repression system based on the interaction of piRNA with PIWI proteins (P-element-induced wimpy testis) (58). Later, similar complexes have been described in the mouse and human germ lines where they play a role in the regulation of L1 elements (58).

PIWI proteins belong to the Argonaute superfamily (58). Three PIWI proteins have been identified in mice (four in humans), *MILI*, *MIWI*, and *MIWI2*, and all are mainly germ line restricted (59). In 2006, thanks to the development of deep-sequencing techniques, research groups identified specific populations of small RNAs binding to PIWI proteins: the piRNAs (58). These piRNAs are mainly generated from dedicated loci, called piRNA clusters (60), consisting of 20–90 kb genomic regions. Around 20% of them matched repetitive elements (61), thereby allowing to distinguish two kinds of piRNAs: repeat associated and non-repeat associated piRNAs. Among the repeat associated piRNAs, which are mainly derived from retrotransposons, we can set apart primary and secondary piRNAs. The one belonging to the piRNA category is length and sequence specific: primary piRNAs are about 26 nucleotides (nt) in length, have a uracil nucleotide at position 1, and correspond to sense-strand transcripts, which predominantly associate with MILI protein. Secondary piRNAs are 31 nt in length, have an adenine nucleotide at position 10, and correspond to antisense strand, bound by MIWI2 protein (58). However, this pathway is still not entirely understood and several open questions need to be addressed. In particular, the initiation of the process is not yet fully understood as well as the loading of the

piRNAs into the PIWI proteins. However, recent studies performed in *Drosophila* have shown that a complex composed of Rhino, Deadlock, and Cutoff (RDC) suppresses piRNA cluster transcript splicing and drives piRNA biogenesis (62, 63). Other evidences show that *MIWI2* loading is strictly dependent of *MILI* SLICING activity (64), and that primary and secondary piRNAs are located in distinct cytoplasmic compartments. Primary piRNAs loaded on MILI supported by the Tudor domain-containing TDRD1 protein are located in specific organelles named pi-bodies (65), whereas secondary piRNAs loaded on MIWI2 supported by TDRD9 protein are located in distinct compartments named piP-bodies (66). Both complexes have been described to be involved in L1 regulation in the germ line (67, 68). *Mili* and *Miwi2* mouse mutant males are sterile, and L1 transcripts accumulate strongly in their germ line (61, 69), phenotypically mimicking the *Dnmt3L* mouse mutant phenotype (70). These similarities suggest that the piRNA pathway plays a role in the control of L1 transcriptional output by promoting *de novo* methylation (64). Finally, a study has shown that in *Drosophila*, Piwi exhibits a slicer activity: it can bind single-strand RNA and cleave the corresponding RNA target (71). This implies that the piRNA machinery could also be involved in the post-transcriptional silencing of L1.

Few reports have investigated the role of PIWI-piRNA pathways in L1 regulation in humans. The Larriba laboratory showed that hypermethylation-associated silencing of PIWI2 and TDRD1 is correlated with the hypomethylation of L1 and linked with spermatogenic disorders. This suggests that in humans, PIWI-piRNA pathways could also target L1 elements and that the reactivation of such TEs could contribute to spermatogenic failures (72).

Regulation in the mature gametes

From 13.5 to 16.5 dpc, the genome of male germ cells undergoes a wave of remethylation (Figure 3A) (43). Moreover, L1 elements of epididymal sperms from adult testis have been shown to be fully methylated (73). In contrast, female germ cells are never fully methylated (73, 74), whereas secondary oocytes (ovulated but unfertilized) show an intermediate level of L1 methylation (around 42%) (75) (Figure 3A). Therefore, L1 are expressed in mature oocytes (76), which implies the presence of L1 regulatory mechanisms at the oocyte stage other than methylation.

A previous study revealed that exogenously introduced target RNAs containing L1 sequences were specifically degraded in oocytes by an RNA interference (RNAi)-dependent mechanism (77), suggesting that in mouse

oocytes, L1 could be repressed through the RNAi pathway (described in ‘RNAi pathways in mammals’). This idea was strengthened by the fact that oocytes produce small interfering RNAs (siRNAs) derived from L1 sequences (78). However, *Dicer*-deficient mutant oocytes did not show any L1 accumulation (78). This last observation could be explained by the identification of specific mouse oocyte piRNAs (78) (see ‘Role of the piRNA pathway’), which could also regulate L1 elements in oocytes.

Moreover, mammal oocytes are blocked at the diplotene stage of meiosis I from the fetal stage until sexual puberty. This long meiotic arrest could be an innate barrier to the accomplishment of the full L1 retrotransposition cycle (Figure 1B). Interestingly, during development, oocytes also undergo extensive apoptosis, leading to a massive oocytes loss; this process is called fetal oocyte attrition (FOA) and is still poorly understood. A recent study, however, revealed that enhancement of L1 expression is involved in the reduction in the number of fetal oocytes at birth and that FOA can be modulated by controlling L1 activity (79). It is proposed that FOA corresponds to a select mechanism by which only oocytes with low L1 activity are conserved (79).

Regulation in preimplantation embryo

Shortly after implantation, at around 6.5 dpc, the mouse embryo undergoes a wave of *de novo* methylation, establishing a genome-wide hypermethylation pattern (74, 75). However, during the development of preimplantation embryos, from the zygote to the blastocyst stage, a structure formed around 2.5 days after fertilization, L1 elements are subject to a first wave of demethylation (Figure 3A). Indeed, a dramatic reduction of L1 methylation level is observed during the sperm-to-zygote transition and then methylation progressively decreases to reach 13–23% at the blastocyst stage (43, 74). This can lead to L1 transcript accumulation in the blastocyst and therefore to retrotransposition events (80). Moreover, other evidences have shown that L1 RNA transcribed in male or female germ line could remain competent for integration in the early mouse embryo after being carried over by the gametes through fertilization in mice (81) and in humans (82). These data imply that the preimplantation phase can be considered as an aperture favorable for L1 invasion.

DNA methylation during preimplantation

In mammals, DNA methylation occurs at the 5-position of cytosine (5mC) and is known to be essential for

development and many other biological processes, as a transcriptional regulation mechanism and for the maintenance of genome stability (47, 48). Until 5-hydroxymethylcytosine (5hmC) was discovered in mammals in 2009, it was the only known DNA epigenetic mark (83, 84). It has been shown that the ten-eleven translocation (TET) protein family is responsible for the conversion from 5mC into 5hmC through oxidation (85). This suggests that 5hmC may serve as an intermediate of DNA demethylation. As 5hmC is not recognized by the maintenance DNA-methyltransferase DNMT1 during DNA replication (86), it is proposed that the conversion of 5mC to 5hmC leads to the restriction of DNA methylation patterns through passive DNA demethylation during cell division. Further studies showed that TET proteins could oxidize 5hmC to produce 5-formylsytosine (5fC) and 5-carboxylcytosine (5caC) that can be removed from the genome by a thymidine-DNA glycosylase (87, 88). Evidences suggest that 5fC and 5caC might serve as intermediates in an active DNA demethylation process, as in *Tdg* gene knockout mouse, leading to increased DNA methylation at certain genomic loci (89).

After fertilization, the level of 5hmC of the paternal genome increases strongly until the end of the zygote stage. Subsequently, it decreases to reach its basic level (Figure 3B). Later, during the blastocyst stage, the 5hmC level of both paternal and maternal genomes highly increases (Figure 3B) (47). Several lines of evidence support the notion that TET3 drives the 5mC oxidation during preimplantation and also DNA demethylation: (i) during the same time window (from zygote to blastocyst stage), an increase of the 5hmC level and a decrease of the global L1 methylation level are observed (Figure 3A); (ii) TET3 is highly expressed at the zygote stage (90); and (iii) Tet3 siRNA-mediated knockdown abolishes 5mC oxidation (91). Nonetheless, the role of 5hmC in the regulation of TEs still needs to be further investigated.

However, a recent report describes a gonad-specific expression gene (GSE) to be present specifically in germ cells and in preimplantation embryos. This protein is localized in the nuclei of cells from the zygote to the blastocyst stage (92). Zygote GSE knockdown mediated by antisense RNA results in an increase in LINE-1 methylation level and reduction of 5mC to 5hmC on L1 (92), suggesting an active DNA demethylation role for GSE during zygote development and a role of 5hmC in LINE-1 regulation.

Finally, two recent articles described the methylation profile during human and mouse early development (93, 94). These studies reveal that the global methylation reprogramming after fertilization is globally conserved.

Nonetheless, L1 with different evolutionary ages show different demethylation patterns: young L1 elements are more resistant to demethylation than their older counterparts in human embryos (93).

RNAi pathways in mammals

In several organisms, RNAi pathways are used as a post-transcriptional mechanism for TE repression (95). Three major classes of endogenous small RNAs have been first identified in mammals: microRNAs (miRNAs), piRNAs, and endogenous siRNAs (endo-siRNAs) (58). While miRNAs and siRNAs are loaded into the AGO effector proteins, piRNAs interact only with the PIWI proteins, expressed mainly in germ cells (see 'Role of the piRNA pathway') (58).

Dicer promotes processive cleavage of double-stranded RNAs into endo-siRNAs (96) that are loaded into AGO2 to mediate endonucleolytic cleavage of their target transcript. miRNAs have been identified as important post-transcriptional regulators of gene expression (97). Most animal miRNAs are transcribed in the nucleus by RNA polymerase II as stem-loop long primary transcripts (pri-miRNA). Following transcription, they are processed sequentially in the nucleus and cytoplasm by a complex of RNase-III endonucleases: DROSHA and DICER. Specifically, DROSHA and its partner DGCR8, which represent what we call 'the microprocessor complex', process the pri-miRNA transcript to a 70–93 nt stem-loop precursor RNA (pre-miRNA). Subsequently, they are delivered to the cytoplasm by Exportin 5, where they are cleaved by a non-processive DICER, which forms a complex with TRBP, to produce an ~22 nt miRNA-3p:miRNA-5p duplex (canonical pathway). Once in the cytoplasm, one of the duplex strands (miRNA-3p or miRNA-5p) is preferentially incorporated into the RNA-induced silencing complex (RISC) in association with an AGO family member. Within the RISC-AGO entity, the miRNA guides the complex to its RNA target, thereby mediating its repression. In animals, miRNAs control gene expression by binding to the 3'UTR of their target genes through Watson-Crick base pairing between the target and the 5'-end of the miRNAs: the 'seed' sequence (nt 2–8). Furthermore, the Microprocessor is directly involved in the regulation of L1 elements in humans (described in 'Regulation in the blastocyst') (38).

Finally, recent technical developments have enabled the identification of a multitude of novel types of small RNA molecules that do not fit into the well-established classes (35, 98). Moreover, RNAi pathways expand beyond the post-transcriptional regulation, by having a role in the

maintenance of cellular integrity (35) and by acting as an antiviral defense mechanism (99).

A shift of the small RNA balance during preimplantation

After fertilization, the first event of gene expression is called zygotic gene activation. It takes place at the two-cell stage in mice and at the four-cell stage in humans, and has been involved in qualitative and quantitative changes in coding and non-coding gene expression (100, 101). In zebrafish, an miRNA (miR-430) has been involved in the deadenylation and the clearance of maternal mRNAs (102). This result and the development of sequencing technologies have encouraged researchers to profile small RNAs during mouse early development from mouse unfertilized (metaphase II: MII) oocytes, 8–16-cell stage embryos to the blastocyst stage (103). This study revealed a switch in small RNA populations from siRNA and piRNA derived from retrotransposons to the transcription of miRNAs (103).

How and why this change in small RNA populations occurs still needs to be further investigated. Recently, it has been shown that the processivity of DICER, the key enzyme in siRNA and miRNA pathways, decreases during early development (104), which could explain the switch between siRNA to miRNA populations. The decrease of the piRNA population correlates with the first wave of demethylation of the genome (see ‘DNA methylation during preimplantation’), suggesting a cooperative role of piRNA and methylation pathways.

L1 mRNAs have been recently monitored by RNA-seq and RT-qPCR from the two-cell stage to the blastocyst stage (105). This work revealed that L1 are reactivated after fertilization and intensively transcribed at the two-cell stage embryo; however, their expression strongly decreases until the eight-cell stage and is maintained at a low level until the blastocyst stage, suggesting the existence of an active and fast repression of L1 elements during preimplantation. However, little is known about the regulation of L1 during early development. The expression profiles of the *Piwi* family genes indicated that MILI protein is transiently expressed at the eight-cell stage (103), suggesting the possibility of a regulation through the piRNA pathway. In contrast, experiments performed on fertilized 1- and 8–16-cell embryos, involving the introduction of GFP RNAs carrying L1 sequences and the monitoring of the target RNA degradation, revealed a specific degradation of the target RNAs at both stages: 1-cell and 8–16-cell embryo (103). Finally, maternally derived L1 small RNAs also appeared to be active to at least the 8–16-cell stage (103).

Regulation in the blastocyst

The hypomethylated status of L1 elements (75), the lack of piRNAs (103), and the ability of engineered L1 to integrate during embryogenesis (81) turn the blastocyst stage into a critical window for L1 regulation. mESCs are pluripotent stem cells derived from the inner cell mass of blastocysts. They can be maintained undifferentiated under controlled culture conditions or induced into the three primary germ layers, therefore representing a useful model to mimic *in vitro* early development.

Profiling of small RNAs from mESCs revealed the presence of miRNAs and some endogenous siRNAs derived from repeated elements (106, 107), but not piRNAs. These observations bring them up as a suitable model to study the mechanism that might regulate L1 expression and retrotransposition during the last step of preimplantation development.

Both human and mouse L1 5'UTR contain sense and antisense promoters. This bidirectional transcription of L1 5'UTR has the potential to generate specific double-stranded RNAs, which could be a perfect substrate for Dicer (34–36). Moreover, experiments in human cultured cells have shown that the RNA transcribed from the 5'UTR antisense promoter induces post-transcriptional mRNA degradation of the 5'UTR sense derived transcript through RNAi, suggesting a role of DICER in the regulation of L1 elements in human cells (108). Experiments performed in *Dicer* knockout mESCs showed an upregulation of mRNAs and ORF1 protein derived from L1, hypomethylation of the 5'UTR, and a gain of L1 copy number in the genome (35). These observations, combined with the fact that *Ago* knockout mESCs also shows an upregulation of L1 transcripts and an increase of L1 copy number (35), demonstrate that the siRNA pathway is involved in the L1 regulation in mESCs and probably also at the blastocyst stage of the mouse embryos.

Other evidences suggest the involvement of another actor of the RNAi pathway in the regulation of human L1 element: the microprocessor. It has been shown that the Microprocessor is able to bind L1-derived RNAs, to regulate the level of L1 mRNAs and ORF1 protein, and to cleave *in vitro* the 5'UTR of L1 mRNAs (38). In the proposed model, the microprocessor restricts L1 retrotransposition at a post-transcriptional level by binding L1 mRNAs within the nucleus and cleaving hairpin structures contained in the L1 5'UTR (Figure 2C). This process has the potential to destabilize L1 transcripts and decrease the production of L1 proteins, and finally to reduce retrotransposition (38). A similar process has already been described for regulation at the human immunodeficiency virus (HIV)

type 1 promoter (109), where it was reported that the microprocessor can cleave the nascent RNA and generate an uncapped transcript. This uncapped RNA serves as a signal for the recruitment of a termination factor that degrades the ongoing transcript, leading to the termination of transcription (109).

Recently, a study from the Trono laboratory demonstrated that a specific subset of L1 elements is repressed by the KAP1 protein, a KRAB-containing zinc finger protein cofactor, in human ESCs. Moreover, the mutation of this gene in mESCs reactivates the expression of old L1 elements (45). The authors proposed an evolutionary model in which newly emerged L1 lineages are first repressed by DNA methylation before being taken over by a KAP1 protein-mediated silencing process (45). This model is consistent with the observation made in human early embryos by the Qiao laboratory (93), which described that young L1 elements are more resistant to demethylation than their older counterparts (see ‘DNA methylation during preimplantation’).

Finally, a recent report from the Junewein laboratory demonstrated that heterochromatin is required to restrict the aberrant expression of TEs and that Suvar39 histone methyltransferase is important for the silencing of L1 type A in mESCs (110).

However, it must be remembered that mESCs and human ESCs are not equivalent: they use different signaling pathways to maintain their pluripotency and diverge at the epigenetic state (111). Data obtained on LINE-1 elements in mESCs are not necessarily applicable to human L1 and *vice versa*. Moreover, several reports based on transcriptome analyses proposed that human ESCs are more closely related to post-implantation mouse epiblast stem cells than to mESCs (111, 112).

Regulation in somatic cells

DNA methylation patterns are established during embryonic development by *de novo* methylated enzymes, DNMT3A and DNMT3B (see ‘Role of DNA methylation’). In mammals, this methylation at 5mC is mainly maintained by the maintenance DNA methyltransferase DNMT1 (48). Moreover, this epigenetic modification can inactivate genes by direct exclusion of the transcriptional machinery from methylated promoter DNA. It acts by directly impeding the binding of transcriptional factors to their target sites, and altering chromatin structure through histone modification and nucleosome occupancy within the promoter regions of genes (48).

The implication of DNA methylation in regulating L1 expression originates from multiple studies. In various

cell lines, the expression of L1 full-length transcripts is correlated with the differential methylation level of the promoter region (47, 48). In somatic cells, few or no L1 proteins are expressed, and most L1 elements are fully methylated. On the contrary, a full-length RNA signal becomes detectable if methylation is inhibited, and the level of 5'UTR-containing transcripts is increased in cultured cells (113, 114). However, DNA methylation is not the only L1 regulatory mechanism present in somatic cells, and in the following section we will also focus on the role of RNA editases in L1 element repression.

L1 methylation in somatic tissues

Cell-based reporter assays performed in human cells showed that L1 promoter activity is significantly inhibited by CpG methylation and that only a subset of these CpG sites affects L1 promoter activity in an *in vitro* transcription assay when mutated (115). Methylated CpG sites are recognized by members of the methyl-CpG-binding domain (MBD) protein family, which in turn recruit histone deacetylases and generate transcriptionally inactive chromatin structures (116). Reporter assays have shown that methyl-CpG-binding protein 2 (MeCP2), the founding member of the MBD family, can bind to methylated L1 5'UTR and represses its transcription (117, 118).

The characterization of L1 5'UTR made it possible to specifically investigate the methylation status of L1 promoters, and a large number of CpG sites in L1 5'UTR have been surveyed by bisulfite sequencing (119). In general, the methylation patterns of the 5'UTR during development are consistent with those of the body of L1. Finally, it has been shown that the body of endogenous L1 is fully methylated in somatic tissues such as the brain, kidney, liver, and spleen (75).

Role of the RNA editases

RNA editing is a molecular process that allows the modification of the information content in an RNA molecule by switching nucleotide sequences, usually cytidine into uridine (C to U) or adenosine into inosine (A to I) (120). RNA-dependent deaminases are the enzymes responsible for the base deamination process. By altering the RNA, and consequently the amino acid sequence of encoded proteins, RNA editing plays an important role in expanding the genome capacity. Mutant mouse models for such enzymes show immunity issues and susceptibility to viral infection (121), suggesting that RNA-editing proteins are

involved in innate immune responses against infectious RNA viruses. As RNA editases can act against invading RNA particles, they can also be potential inhibitors of TEs.

Among them, we can highlight the APOBEC protein family that catalyzes the deamination of cytosine residues into uracils. It has been shown that one member, APOBEC3G, can reduce the replication of HIV by inducing uracil mutation accumulation on the retroviral complementary strand, and thus inactivating the newly integrated copy (122). Additionally, retrotransposition assay experiments have highlighted APOBEC3A and APOBEC3B as potent inhibitors of retrotransposons (123). Recent work from the Moran laboratory demonstrated that APOBEC3A can inhibit L1 retrotransposition by deaminating transiently exposed single-stranded DNA that arises during new L1 integration events (124). Similarly, the activation-induced deaminase (AID) protein is also able to inhibit L1 retrotransposition, through a mechanism independent from deamination (125) and not yet understood. It has been hypothesized that this family mediates cytoplasmic sequestration of L1 RNA and/or L1-encoded proteins, or inhibits the activity of L1 ORFs (126).

The special case of the neuronal progenitor cells

The neural stem cells residing in the neurogenic regions of the brain have the abilities to remain multipotent, replicate, and differentiate into glial progenitors or neuronal progenitor cells (NPCs). L1 promoters are highly methylated in neural stem cells, and MeCP2 is recruited for their repression (117, 118). However, it has been observed that not only L1 expression increases upon the NPC differentiation (127) but also that L1 can retrotranspose in human and mouse NPCs *in vitro* (127, 128). These L1 integration events lead to an alteration of the expression of neighboring genes by promoter enhancement and epigenetic silencing (127), which could affect neuronal fate and function. The details of these retrotransposition events and their possible consequences have been recently reviewed (129). This unique ability of L1 to retrotranspose in NPCs could be explained by its repression mechanism involving the TF sex determining-region Y-box 2 (Sox2). It has been shown that the L1 promoter contains several TF binding sites, including a Sox2 site (130). This key factor is involved in the maintenance of the multipotency and the proliferative state of neural stem cells (131). However, Sox2 is also involved in the inhibition of L1 expression in neural stem cells by forming a repressor complex and associating with the L1 promoter (127). When neural stem cells commit to the neuronal lineage, Sox2 expression decreases, repressor

complexes dissociate, and L1 promoter methylation level drops (127), thus explaining the high expression of L1 in NPCs. Interestingly, the promoter region of a major TF promoting neurogenesis, NeuroD1, harbors a Sox2 binding domain similar to the L1 5'UTR (129). Hence, expression of NeuroD1 and L1 could follow the same pattern during neurogenesis, explaining the high L1 expression during neuronal differentiation. Finally, a study has shown enrichment of 5hmC in Purkinje neurons (83). As we have previously explained (see 'DNA methylation during pre-implantation') that this epigenetic mark is related to DNA demethylation and therefore to the relaxation of L1 repression, this observation could explain the higher sensitivity of neuronal precursors to L1 mobilization.

Other LINE-1 regulators described in somatic cells

Numerous studies performed on human cell lines have recently revealed new putative L1 regulators (132). Here, we present some newly described examples in the literature. An RNA helicase, the MOV10 protein, has been shown to be able to inhibit retrotransposition (133). MOV10 is able to associate with the L1 RNP particle and reduces L1 retrotransposition activity when it is overexpressed in human cell lines (133). The role of MOV10 as a potential L1 inhibitor is strengthened by the significant increase of L1 retrotransposition events in MOV10-depleted cells. However, the inhibitory mechanism involving MOV10 is still unclear. MOV10 has been shown to interact with the RISC, indicating that RNAi pathways could be involved in L1 regulation in human cell lines. In mice, MOV10-like-1, the testis specific paralogue of MOV10, is able to bind the piRNA-associated proteins MILI and MIWI (134). Finally, in the testis of *Mov10-like-1* knockout mice, an upregulation of L1 mRNA has been observed (134), suggesting a role of piRNA pathway in L1 repression.

The work of the Martin laboratory has highlighted other classes of L1 repressors such as the heterogeneous nuclear RNPs (hnRNPs), which are composed of four members: R, Q, L, and nucleolin (NLC) (135). In their study, the authors performed an *hnRNPL* knockdown in human cell lines and observed a significant increase of ORF1p level, L1 mRNA level, and a 10-fold increase of L1 retrotransposition events, implying a role of hnRNPL as a potent negative regulator of L1 in human cell lines (135). On the contrary, depletion of NLC leads to a decrease of L1 retrotransposition rate by about 10-fold, designating NLC as a positive actor in the L1 retrotransposition process (135). As hnRNPL plays multiple roles in RNA metabolism and no effect on the L1 mRNA processing was observed

in the depleted cells, the research team hypothesized that hnRNPL acts as defense factor by decreasing the L1 mRNA and therefore the level of RNP available for the final step of the retrotransposition cycle.

Furthermore, in human cells, Boeke's team showed that two poly(A) binding proteins (PABPs), PABPN1 and PABPC1, can affect the L1 retrotransposition activity (136). They proved that both of them are associated with L1 RNP complexes, probably through the ability of both proteins to bind the L1 mRNA. Moreover, knockdown experiments of the corresponding mRNA result in significant reduction of L1 retrotransposition rate. By opposition, in PAIP2 knockdown cells, an inhibitor of PABPC1, a 2-fold L1 retrotransposition increase has been observed (136). Taken together, these results suggest that PABPC1 is essential for the formation of the RNP, and any depletion of PABPC1 can lead to defect in the RNP formation and in the achievement of a retrotransposition cycle (136). As it is known that PABPC1 can shuttle between the cytoplasm and the nucleus, the authors assumed that PABPC1 might provide a direct trafficking function for the RNP nuclear translocation, an essential step to complete a retrotransposition cycle (136).

Finally, another study proposes the 3' repair exonuclease Trex-1 as a potent regulator of L1 (137). Trex-1 is a negative regulator of the interferon-stimulatory DNA pathway, a cell-intrinsic antiviral response activated by cytosolic DNA detection. An accumulation of single-stranded DNA fragment derived from endogenous retroelements is observed in Trex-1-deficient cells, and it has been shown that Trex-1 can metabolize specifically reverse-transcribed DNA of endogenous retroelements (137). These evidences suggest a connection between retroelements and the interference DNA signaling (IDS) pathway. However, the mechanism of metabolization remains unclear.

LINE-1 in cancers

The hypomethylation status of L1 5'UTR in malignant cells has been correlated with the L1 expression levels, and numerous cancers showed distinct L1 activation. What is known about L1 activity during oncogenesis and tumor progression has already been reviewed intensively recently (138, 139), and we refer to these articles for more details. However, even if a reactivation of L1 is observed in such tissues, a crucial question remains elusive: is L1 mobilization a cause or a consequence of cancer progression?

In this line of idea, a recent study analyzed 290 cancer samples from 244 patients across 12 tumor types and

revealed that most retrotransposition events are harmless, arguing in favor of the fact that a huge amount of L1 retrotransposition will be a consequence of tumorigenesis (140).

Challenges in analyzing the epigenetic status of repeat elements

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) has become a gold standard technique to genome-wide profiling of chromatin signatures, including DNA-binding proteins and histone modifications. Up to date, vast amounts of ChIP-seq data have been generated, focusing mainly on the epigenetic patterns of genes, promoters, and enhancers, but not TEs (141), thereby leaving this area largely unexplored (142).

Usually, ChIP-seq data analysis can be divided into the following steps: quality control, read mapping, peak calling, peak annotation, and motif identification (143). For each step, algorithms have been well developed and a series of bioinformatic software are available (142). However, the main challenge in analyzing ChIP-seq data on TEs is still how to allocate multiple-hit reads (MHRs; i.e., reads that could be mapped to several equally good positions). The small size of next-generation sequencing reads (36–50 bp with the first generation of deep-sequencing machines) and the huge amount of repeat elements in the genome are therefore the main issues. In this part, we discuss two different strategies already used to annotate reads on L1 and compare their pros and cons.

Mapping on consensus sequences

One solution to avoid, but not directly solve, this problem is to identify the peaks at the level of the consensus sequence for each repeat type or family, and not on the genome sequences (105, 144, 145). For example, instead of discarding all MHRs, Day et al. (144) kept the reads that could be mapped to the same or similar repeat type using a phylogenetic approach. By implementing this method with histone modification ChIP-seq dataset from mouse cell lines, they achieved a 10-fold increase of read usage for repeat elements. Indeed, the authors identified distinct histone modification patterns between endogenous retrovirus (ERV)-K and ERV1 subfamilies, in a subtype of LTR retrotransposons. It will now be

interesting to implement this method for other kinds of repeats, to investigate their epigenetic pattern. However, as the enrichment estimation is based on a consensus repeat sequence, other approaches are still needed to investigate the epigenetic pattern in every individual repeat element in the genome.

Using computational approaches to rescue MHRs

A more direct way to address MHRs is to allocate them into potential regions according to known information, e.g., the number of unique-hit reads (UHRs). A similar strategy is currently used for isoform quantification from RNA-seq data and could be addressed using similar algorithms implemented in CuffLink (146) and RSEM (147). This type of methods could rescue all MHRs, which are usually filtered out in ChIP-seq data analysis. For example, Chung et al. (148) conducted an expectation-maximization algorithm to allocate the MHRs to each potential region according to the count of UHRs. By implementing this algorithm for STAT1 and GATA1 ChIP-seq libraries in human and mouse cells, they identified up to 35% novel binding sites, most of them being located in repeat elements and segmental duplications (148). More recently, Wang et al. (149) also developed an empirical formula to distribute MHRs according to two factors: the distance of MHR to its nearest peak identified by UHR and the enrichment score of the peak. They implemented this algorithm to several datasets, including DNA methylation, histone modifications, and TF binding, and identified up to 60% new peaks (138).

Strategies remain imperfect

However, all these methods rely on assumptions, which could be not true. For example, the distribution of UHRs may not reflect the true binding preference of a TF and thus will mislead the allocation of MHRs. Also, an empirical formula derived from one experimental design may not fit another setup. The current approach in the field of TEs to address this problem is to increase the read size (100–200 bp) and use a paired-end library, or even develop new experimental protocols that could specifically capture repeat elements (110, 138). Besides, special attention should be paid when mapping reads to species with only draft genome (e.g., sloth and tarsier), where the annotation of repeat elements is far from comprehensive.

Expert opinion

In this review, we tried to provide an overview about what is known concerning the regulation of the most represented TE in the mouse and human genomes. Taking into account the high mutagenic potential of L1, due to their ability to ‘jump’, specific repressive mechanisms have been established in mammals to prevent invasion by these mobile elements during the life cycle. First, we compared the L1 elements from the two major models used for their study: the mouse and the human L1. We showed that even if they share common features, they also differ in their mobilisation activity and at the structure level, particularly of the 5'UTR, a key region for the expression and mobilization of L1. The differences observed between the predicted secondary structures of mouse and human L1 5'UTR (Figure 2C) could involve the existence of different L1 regulatory mechanisms between this two species. This implies that results derived from studies performed with human L1 constructs in mouse or *vice versa* should be considered with more caution.

Subsequently, we presented the major L1 defense mechanisms currently described in the literature. We showed that over the life cycle of a mammalian individual, a range of mechanisms regulate the remaining L1 active copies present in the genome. In somatic cells, studies revealed a potential role of RNA editases in the repression of L1. However, DNA methylation seems to be the main regulatory mechanism in somatic cells to ensure transcriptional repression of these mobile elements. During embryonic development, mammals undergo two waves of demethylation, thus establishing windows of opportunity for L1 expression and retrotransposition. The first wave starts at the zygote stage and continues until the blastocyst stage (Figure 3A). It has been shown that L1 are reactivated and expressed just after fertilization; however, this expression rapidly decreases as development progresses. Little is known about the L1 regulation during early embryonic development. However, evidences indicate that RNAi pathways are involved. In addition, further results suggest that piRNA could be involved before the blastocyst stage. During the blastocyst stage, owing to the lack of piRNAs, other models have been proposed to explain L1 regulation involving the Microprocessor complex, which could directly bind and cleave human L1 transcripts. Alternatively, DICER protein and the siRNA pathway are involved in mice. The previous model is supported by the presence of sense and antisense promoters in L1 5'UTR of both humans and mice, which have the potential to generate specific double-stranded RNAs that are suitable substrates for DICER.

Later during development, the second wave of demethylation occurs during the migration of the PGC precursor cells (Figure 3A). During this phase, the piRNA pathway seems to be the major L1 regulatory mechanism. It is suggested that the PIWI proteins, by loading L1 piRNAs, could not only promote specific *de novo* methylation but also cleave the corresponding target. This implicates that the piRNA pathway could regulate L1 at two levels: transcriptional and post-transcriptional. However, how the piRNAs are loaded into the PIWI proteins or how the PIWI proteins could recruit the *de novo* methylation machinery are still not understood and remain the major questions in this field. At the level of the mature gamete, the piRNA pathway has been proposed to maintain L1 repression in oocytes considering the low level of methylation of its genome compared with the male gamete. Recently, a publication has highlighted the FOA as a possible process of selecting suitable oocytes by eliminating the female gamete possessing a high level of L1 activity.

Finally, we discussed the challenge in studying the epigenetic mechanisms of such repeat elements. Because of the small size of reads generated from ChIP-seq experiments, accurate mapping is not possible, although several state-of-art computational methods have been developed. Compromises have to be made in this situation, either by mapping reads to consensus sequence, or trying to estimate the MHRs under some assumptions.

Outlook

According to the remaining questions raised in the ‘Expert opinion’ section, future studies should focus on the mechanisms driving the piRNA pathway. In particular, the studies should investigate the generation of the piRNAs, the loading of these piRNAs into the PIWI proteins, and the possible recruitment of the *de novo* methylation machinery by the PIWI proteins.

The regulation of L1 during preimplantation development should also be further investigated. The development of new genetic screens in mammals, thanks to the establishment of the haploid stem cell system, could be a system to discover new actors in L1 regulation. Moreover, the revolution at the genome engineering level with the discovery and the development of the CRISPR/Cas9 system in mammalian cells should be used to manipulate the L1 expression level in all cell types, both *in vitro* and *in vivo*.

Furthermore, it has also been shown that mobile elements are responsive to stress; however, this area is

understudied and should be addressed further in the future. A good example is the effects of hypoxic stress, which are frequently associated with cancer progression and have been recently linked to an overrepresentation of SINEs in blind mole rats (150).

Highlights

- L1 elements are regulated by a large range of mechanisms during the mammalian life cycle. These mechanisms are supposed to overlap to reinforce L1 repression. However, the question of recognition still remains: how are L1 elements initially recognized and triggered?
- Similarly, the question concerning the origin of L1 remains elusive. Consequently, the questions linked to the establishment of host defense mechanisms, the evolutionary strategies used by mobile elements to get past these mechanisms, as well as the arm race between jumping elements and the host organism are still unsolved and need to be investigated.
- The question of the biological significance also subsists. Do L1 have a biological role or can we just consider them as parasites tempting to expand in the host genome? Some evidences suggest that L1 are required for cell differentiation and early development.
- L1 reactivation has been observed in cancer cells or tissues. However, it is still not possible to determine if this reactivation is the cause or a consequence of the cancer.
- Finally, the factors making some L1 prone to activation and reactivation need to be investigated more deeply. Stem cells and NPCs have specific epigenetic status that could explain L1 activation. Indeed, studies of different pluripotent cell lineages could be particularly interesting to better understand L1 regulation.

Acknowledgments: We thank Dr. Tobias Beyer and the Ciaudo laboratory for the critical reading of the manuscript and for fruitful discussions. This project was supported by a core grant from ETH-Z (supported by Roche) and a grant from Novartis Stiftung für Biologisch-Medizinische Forschung (13B063). M.B. is supported by a PhD fellowship from the ETH-Z foundation (ETH-21 13-1). We apologize if relevant publications were not cited due to space constraints.

Conflict of interest statement: The authors declare no conflict of interest.

References

- McClintock B. [The origin and behavior of mutable loci in maize](#). *Proc Natl Acad Sci U S A* 1950; 36: 344–55.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrum J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissole SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickinson M, Grimwood J, Cox DR, Olson MV, Kaul R, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran J V, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, Szustakowski J. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860–921.
- Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyas E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigó R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrum J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Raymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang S-P, Zdobnov EM, Zody MC, Lander ES. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002; 420: 520–62.
- Finnegan DJ. [Eukaryotic transposable elements and genome evolution](#). *Trends Genet* 1989; 5: 103–7.
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH. [A unified classification system for eukaryotic transposable elements](#). *Nat Rev Genet* 2007; 8: 973–82.
- Engels WR, Johnson-Schlitz DM, Eggleston WB, Svedt J. [High-frequency P element loss in Drosophila is homolog dependent](#). *Cell* 1990; 62: 515–25.
- Kazazian Jr HH, Scott A. “Copy and paste” transposable elements in the human genome. *J Clin Invest* 1993; 91: 1859–60.
- Kolosha VO, Martin SL. In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc Natl Acad Sci U S A* 1997; 94: 10155–60.
- Mathias S, Scott A, Kazazian Jr HH, Boeke J, Gabriel A. Reverse transcriptase encoded by a human transposable element. *Science* 1991; 254: 1808–10.
- Feng Q, Moran JV, Kazazian HH, Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 1996; 87: 905–16.
- Doucet J, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, Huira C. Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet* 2010; 6.
- Ostertag EM, Kazazian HHJ. Biology of mammalian L1 retrotransposons. *Annu Rev Genet* 2001; 35: 501–38.
- Ostertag EM, Kazazian HH. Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. *Genome Res* 2001; 11: 2059–65.

14. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A* 2003; 100: 5280–5.
15. Naas TP, DeBerardinis RJ, Moran J V, Ostertag EM, Kingsmore SF, Seldin MF, Hayashizaki Y, Martin SL, Kazazian HH. An actively retrotransposing, novel subfamily of mouse L1 elements. *EMBO J* 1998; 17: 590–7.
16. Goodier JL, Ostertag EM, Du K, Kazazian HH. A novel active L1 retrotransposon subfamily in the mouse. *Genome Res* 2001; 11: 1677–85.
17. Kazazian HH. Mobile elements: drivers of genome evolution. *Science* 2004; 303: 1626–32.
18. Beck CR, Garcia-Perez JL, Badge RM, Moran JV. LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum Genet* 2011; 12: 187–215.
19. Wheelan SJ, Aizawa Y, Han JS, Boeke JD. Gene-breaking: a new paradigm for human retrotransposon-mediated gene evolution. *Genome Res* 2005; 15: 1073–8.
20. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, Boeke JF, Moran JV. Human L1 retrotransposition: cis Preference versus trans complementation. *Mol Cell Biol* 2001; 21: 1429–39.
21. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* 2000; 24: 363–7.
22. Rosser JM, An W. L1 expression and regulation in human and rodents. *Front Biosci* 2012; 2203–25.
23. Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL, Kazazian HH. A mouse model of human L1 retrotransposition. *Nat Genet* 2002; 32: 655–60.
24. Prak ET, Dodson AW, Farkash EA, Kazazian Jr HH. Tracking an embryonic L1 retrotransposition event. *Proc Natl Acad Sci U S A* 2003; 100: 1832–7.
25. Donnell KAO, An W, Schrum CT, Wheelan SJ, Boeke JD. Controlled insertional mutagenesis using a LINE-1 (ORFeus) genetrapp mouse model. *Proc Natl Acad Sci U S A* 2013; 1: 2706–13.
26. Han JS, Boeke JD. A highly active synthetic mammalian retrotransposon. *Nature* 2004; 429: 314–8.
27. DeBerardinis RJ, Goodier JL, Ostertag EM, Kazazian HH. Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. *Nat Genet* 1998; 20: 288–90.
28. Marchetto MCN, Narvaiza I, Denli AM, Benner C, Lazzarini TA, Nathanson JL, Paquola ACM, Desai KN, Herai RH, Weitzman MD, Yeo GW, Muotri AR, Gage FH. Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature* 2013; 503: 525–9.
29. Penzkofer T, Dandekar T, Zemojtel T. L1Base: from functional annotation to prediction of active LINE-1 elements. *Nucleic Acids Res* 2005; 33: D498–500.
30. DeBerardinis RJ, Kazazian HH. Analysis of the promoter from an expanding mouse retrotransposon subfamily. *Genomics* 1999; 56: 317–23.
31. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 2005; 110: 462–7.
32. Loeb DD, Padgett RW, Hardies SC, Shehee WR, Comer MB, Edgell H, Hutchison CA. The sequence of a large L1 element reveals a tandemly repeated 5' end and several features found in retrotransposons. *Mol Cell Biol* 1986; 6: 168.
33. Swergold GD. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol* 1990; 10: 6718–29.
34. Mätlik K, Redik K, Speek M. L1 antisense promoter drives tissue-specific transcription of human genes. *J Biomed Biotech* 2006; 2006: 1–16.
35. Ciaudo C, Jay F, Okamoto I, Chen C-J, Sarazin A, Servant N, Barillot E, Heard E, Voinnet O. RNAi-dependent and independent control of LINE1 accumulation and mobility in mouse embryonic stem cells. *PLoS Genet* 2013; 9: e1003791.
36. Li J, Kannan M, Trivett AL, Liao H, Wu X, Akagi K, Symer DE. An antisense promoter in mouse L1 retrotransposon open reading frame-1 initiates expression of diverse fusion transcripts and limits retrotransposition. *Nucleic Acids Res* 2014; 42: 4546–62.
37. Montoya-Durango DE, Liu Y, Teneng I, Kalbfleisch T, Lacy ME, Steffen MC, Ramos KS. Epigenetic control of mammalian LINE-1 retrotransposon by retinoblastoma proteins. *Mutat Res* 2009; 665: 20–8.
38. Heras SR, Macias S, Plass M, Fernandez N, Cano D, Eyra E, Garcia-Perez JL, Cáceres JF. The Microprocessor controls the activity of mammalian retrotransposons. *Nat Struct Mol Biol* 2013; 20: 1173–81.
39. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 2004; 20: 289–90.
40. Will S, Reiche K, Hofacker IL, Stadler PF, Backofen R. Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. *PLoS Comput Biol* 2007; 3: e65.
41. Lorenz R, Bernhart SH, Höner Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. ViennaRNA Package 2.0. *Algorithms Mol Biol* 2011; 6: 26.
42. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* 1997; 13: 335–40.
43. Lee HJ, Hore TA, Reik W. Reprogramming the methylome: erasing memory and creating diversity. *Cell Stem Cell* 2014; 14: 710–9.
44. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 2007; 8: 272–85.
45. Castro-Diaz N, Ecco G, Coluccio A, Kapopoulou A, Yazdanpanah B, Friedli M, Duc J, Jang SM, Turelli P, Trono D. Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev* 2014; 28: 1397–409.
46. Zamudio N, Bourc'his D. Transposable elements in the mammalian germline: a comfortable niche or a deadly trap? *Heredity (Edinb)* 2010; 105: 92–104.
47. Li E, Zhang Y. DNA methylation in mammals. *Cold Spring Harb Perspect Biol* 2014; 6.
48. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet* 2013; 14: 204–20.
49. Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet* 1998; 20: 116–7.
50. Okano M, Takebayashi S, Okumura K, Li E. Assignment of cytosine-5 DNA methyltransferases Dnmt3a and Dnmt3b to mouse chromosome bands 12A2-A3 and 2H1 by in situ hybridization. *Cytogenet Cell Genet* 1999; 86: 333–4.
51. Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet* 2007; 16: 2272–80.
52. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* 2007; 449: 248–51.

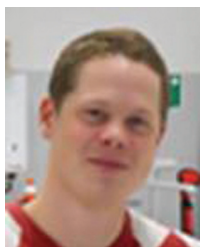
53. Yokomine T, Hata K, Tsudzuki M, Sasaki H. Evolution of the vertebrate DNMT3 gene family: a possible link between existence of DNMT3L and genomic imprinting. *Cytogenet Genome Res* 2006; 113: 75–80.
54. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M, Koseki H. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 2007; 450: 908–12.
55. Huang J, Fan T, Yan Q, Zhu H, Fox S, Issaq HJ, Best L, Gangi L, Munroe D, Muegge K. Lsh, an epigenetic guardian of repetitive elements. *Nucleic Acids Res* 2004; 32: 5019–28.
56. Dunican DS, Cruickshanks HA, Suzuki M, Semple CA, Davey T, Arceci RJ, Greally J, Adams IR, Meehan RR. Lsh regulates LTR retrotransposon repression independently of Dnmt3b function. *Genome Biol* 2013; 14: R146.
57. Pezic D, Manakov SA, Sachidanandam R, Aravin AA. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev* 2014; 28: 1410–28.
58. Siomi MC, Sato K, Pezic D, Aravin AA. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol* 2011; 12: 246–58.
59. Sasaki T, Shiohama A, Minoshima S, Shimizu N. Identification of eight members of the Argonaute family in the human genome. *Genomics* 2003; 82: 323–30.
60. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* 2007; 128: 1089–103.
61. Carmell MA, Girard A, van de Kant HJG, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ, Kant HJG Van De, Bourc D, Rooij DG De. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* 2007; 12: 503–14.
62. Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, Vreven T, Zamore PD, Weng Z, Theurkauf WE. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* 2014; 157: 1353–63.
63. Mohn F, Sienski G, Handler D, Brennecke J. The Rhino-Deadlock-Cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. *Cell* 2014; 157: 1364–79.
64. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* 2008; 31: 785–99.
65. Kojima K, Kuramochi-Miyagawa S, Chuma S, Tanaka T, Nakatsuji N, Kimura T, Nakano T. Associations between PIWI proteins and TDRD1/MTR-1 are critical for integrated subcellular localization in murine male germ cells. *Genes Cells* 2009; 14: 1155–65.
66. Aravin AA, van der Heijden GW, Castañeda J, Vagin VV, Hannon GJ, Bortvin A. Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. *PLoS Genet* 2009; 5: e1000764.
67. Reuter M, Chuma S, Tanaka T, Franz T, Stark A, Pillai RS. Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nat Struct Mol Biol* 2009; 16: 639–46.
68. Shoji M, Tanaka T, Hosokawa M, Reuter M, Stark A, Kato Y, Kondoh G, Okawa K, Chujo T, Suzuki T, Hata K, Martin SL, Noce T, Kuramochi-Miyagawa S, Nakano T, Sasaki H, Pillai RS, Nakatsuji N, Chuma S. The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Dev Cell* 2009; 17: 775–87.
69. Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* 2007; 316: 744–7.
70. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. *Science* 2001; 294: 2536–9.
71. Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes Dev* 2006; 20: 2214–22.
72. Heyn H, Ferreira HJ, Bassas L, Bonache S, Sayols S, Sandoval J, Esteller M, Larriba S. Epigenetic disruption of the PIWI pathway in human spermatogenic disorders. *PLoS One* 2012; 7: e47892.
73. Sanford JP, Clark HJ, Chapman VM, Rossant J. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev* 1987; 1: 1039–46.
74. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 2012; 484: 339–44.
75. Howlett SK, Reik W. Methylation levels of maternal and paternal genomes during preimplantation development. *Development* 1991; 113: 119–27.
76. Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, Knowles BB. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* 2004; 7: 597–606.
77. Watanabe T, Takeda A, Tsukiyama T, Mise K, Okuno T, Sasaki H, Minami N, Imai H. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev* 2006; 20: 1732–43.
78. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, Sasaki H. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 2008; 453: 539–43.
79. Malki S, van der Heijden GW, O'Donnell KA, Martin SL, Bortvin A. A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. *Dev Cell* 2014; 29: 521–33.
80. Packer AI, Manova K, Bachvarova RF. A discrete LINE-1 transcript in mouse blastocysts. *Dev Biol* 1993; 157: 281–3.
81. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev* 2009; 23: 1303–12.
82. Van den Hurk JA, Meij IC, Seleme MDC, Kano H, Nikopoulos K, Hoefsloot LH, Sistermans EA, de Wijs IJ, Mukhopadhyay A, Plomp AS, de Jong PTVM, Kazazian HH, Cremers FPM. L1 retrotransposition can occur early in human embryonic development. *Hum Mol Genet* 2007; 16: 1587–92.
83. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in brain and enriched in Purkinje neurons. *Science* 2009; 324: 929–30.
84. Tahiliani M, Koh KP, Shen Y, Pastor WA, Brudno Y, Agarwal S, Iyer LM, David R, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009; 324: 930–5.

85. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 2010; 466: 1129–33.
86. Valinluck V, Sowers LC. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res* 2007; 67: 946–50.
87. He Y, Li B, Li Z, Liu P, Wang Y, Tang Q. Tet-mediated formation of 5-carboxylcytosine and its excision by TGD in mammalian DNA. *Science* 2011; 333: 1303–7.
88. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 2011; 333: 1300–3.
89. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Coz M, Le, Devarajan K, Wessels A, Soprano D, Abramowitz K, Bartolomei MS, Rambow F, Bassi MR, Fanciulli M, Renner C, Klein-szanto AJ, Kobi D, Davidson I, Alberti C, Larue L. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 2011; 146: 67–79.
90. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W, Walter J. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun* 2011; 2: 241.
91. Gu T-P, Guo F, Yang H, Wu H-P, Xu G-F, Liu W, Xie Z-G, Shi L, He X, Jin S, Iqbal K, Shi YG, Deng Z, Szabó PE, Pfeifer GP, Li J, Xu G-L. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* 2011; 477: 606–10.
92. Hatanaka Y, Shimizu N, Nishikawa S, Tokoro M, Shin S-W, Nishihara T, Amano T, Anzai M, Kato H, Mitani T, Hosoi Y, Kishigami S, Matsumoto K. GSE is a maternal factor involved in active DNA demethylation in zygotes. *PLoS One* 2013; 8: e60205.
93. Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, Yan J, Ren X, Lin S, Li J, Jin X, Shi X, Liu P, Wang X, Wang W, Wei Y, Li X, Guo F, Wu X, Fan X, Yong J, Wen L, Xie SX, Tang F, Qiao J. The DNA methylation landscape of human early embryos. *Nature* 2014; 511: 606–10.
94. Smith ZD, Chan MM, Humm KC, Karnik R, Mekhoubad S, Regev A, Eggan K, Meissner A. DNA methylation dynamics of the human preimplantation embryo. *Nature* 2014; 511: 611–5.
95. Obbard DJ, Gordon KHJ, Buck AH, Jiggins FM. The evolution of RNAi as a defence against viruses and transposable elements. *Philos Trans R Soc Lond B Biol Sci* 2009; 364: 99–115.
96. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 2001; 409: 363–6.
97. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 2008; 9: 102–14.
98. Yamakawa N, Okuyama K, Ogata J, Kanai A, Helwak A, Takamatsu M, Imadome K-I, Takakura K, Chanda B, Kurosaki N, Yamamoto H, Ando K, Matsui H, Inaba T, Kotani A. Novel functional small RNAs are selectively loaded onto mammalian Ago1. *Nucleic Acids Res* 2014; 42: 5289–301.
99. Maillard PV, Ciaudo C, Marchais A, Li Y, Jay F, Ding SW, Voinnet O. Antiviral RNA interference in mammalian cells. *Science* 2013; 342: 235–8.
100. Hamatani T, Carter MG, Sharov AA, Ko MSH. Dynamics of global gene expression changes during mouse preimplantation development. *Dev Cell* 2004; 6: 117–31.
101. Escamilla-Del-Arenal M, da Rocha ST, Heard E. Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Hum Genet* 2011; 130: 307–27.
102. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* 2006; 312: 75–9.
103. Ohnishi Y, Totoki Y, Toyoda A, Watanabe T, Yamamoto Y, Tokunaga K, Sakaki Y, Sasaki H, Hohjoh H. Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. *Nucleic Acids Res* 2010; 38: 5241–51.
104. Flemr M, Malik R, Franke V, Nejepinska J, Sedlacek R, Vlahovicek K, Svoboda P. A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. *Cell* 2013; 155: 807–16.
105. Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, Carninci P, Torres-Padilla M-E. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat Struct Mol Biol* 2013; 20: 332–8.
106. Babiarczyk JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 2008; 22: 2773–85.
107. Ciaudo C, Servant N, Cognat V, Sarazin A, Kieffer E, Viville S, Colot V, Barillot E, Heard E, Voinnet O. Highly dynamic and sex-specific expression of microRNAs during early ES cell differentiation. *PLoS Genet* 2009; 5: e1000620.
108. Yang N, Kazazian HH. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol* 2006; 13: 763–71.
109. Wagschal A, Rousset E, Basavarajiah P, Contreras X, Harwig A, Laurent-Chabalier S, Nakamura M, Chen X, Zhang K, Mezziane O, Boyer F, Parrinello H, Berkhout B, Benkirane M, Kiernan R. Microprocessor, Setx, Xrn2, and Rps6 co-operate to induce premature termination of transcription by RNAPII. *Cell* 2012; 150: 1147–57.
110. Bulut-Karslioglu A, De La Rosa-Velázquez IA, Ramirez F, Barenboim M, Onishi-Seebacher M, Arand J, Galán C, Winter GE, Engist B, Gerle B, O'Sullivan RJ, Martens JHA, Walter J, Manke T, Lachner M, Jenuwein T. Suv39h-dependent h3k9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol Cell* 2014; 277–90.
111. Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RDG. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 2007; 448: 196–9.
112. Brons IGM, Smithers LE, Trotter MWB, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 2007; 448: 191–5.
113. Woodcock DM, Lawler CB, Linsenmeyer ME, Doherty JP, Warren WD. Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. *J Biol Chem* 1997; 272: 7810–6.
114. Tchénio T, Segal-Bendirdjian E, Heidmann T. Generation of processed pseudogenes in murine cells. *EMBO J* 1993; 12: 1487–97.
115. Hata K, Sakaki Y. Identification of critical CpG sites for repression of L1 transcription by DNA methylation. *Gene* 1997; 189: 227–34.

116. Bird A P, Wolffe A P. Methylation-induced repression – belts, braces, and chromatin. *Cell* 1999; 99: 451–4.
117. Yu F, Zingler N, Schumann G, Strätling WH. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucleic Acids Res* 2001; 29: 4493–501.
118. Muotri AR, Marchetto MCN, Coufal NG, Oefner R, Yeo G, Nakashima K, Gage FH. L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 2010; 468: 443–6.
119. Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Reik W. [Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse.](#) *Genesis* 2003; 35: 88–93.
120. Gerber AP, Keller W. [RNA editing by base deamination: more enzymes, more targets, new mysteries.](#) *Trends Biochem Sci* 2001; 26: 376–84.
121. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. [Class switch recombination and hypermutation require activation-induced cytidine deaminase \(AID\), a potential RNA editing enzyme.](#) *Cell* 2000; 102: 553–63.
122. Bishop KN, Holmes RK, Sheehy AM, Davidson NO, Cho S, Malim MH, Louis S. [Cytidine deamination of retroviral DNA by diverse APOBEC proteins.](#) *Curr Biol* 2004; 14: 1392–6.
123. Bogerd HP, Wiegand HL, Doeble BP, Lueders KK, Cullen BR. APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic Acids Res* 2006; 34: 89–95.
124. Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV. APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. *Elife* 2014; 3: e02008.
125. MacDuff DA, Demorest ZL, Harris RS. AID can restrict L1 retrotransposition suggesting a dual role in innate and adaptive immunity. *Nucleic Acids Res* 2009; 37: 1854–67.
126. Beauregard A, Curcio MJ, Belfort M. [The take and given between retrotransposable elements and their host.](#) *Annu Rev Genet* 2008; 42: 587–617.
127. Muotri AR, Chu VT, Marchetto MCN, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 2005; 435: 903–10.
128. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH. L1 retrotransposition in human neural progenitor cells. *Nature* 2009; 460: 1127–31.
129. Thomas CA, Muotri AR. LINE-1: creators of neuronal diversity. *Front Biosci* 2012; 4: 1663–8.
130. Tchénio T, Casella JF, Heidmann T. [Members of the SRY family regulate the human LINE retrotransposons.](#) *Nucleic Acids Res* 2000; 28: 411–5.
131. Bylund M, Andersson E, Novitsch BG, Muhr J. Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nat Neurosci* 2003; 6: 1162–8.
132. Goodier JL, Cheung LE, Kazazian HH. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res* 2013; 41: 7401–19.
133. Goodier JL, Cheung LE, Kazazian HH. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genet* 2012; 8: e1002941.
134. Frost RJA, Hamra FK, Richardson JA, Qi X, Bassel-Duby R, Olson EN. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc Natl Acad Sci U S A* 2010; 107: 11847–52.
135. Peddigari S, Li PW-L, Rabe JL, Martin SL. hnRNPL and nucleolin bind LINE-1 RNA and function as host factors to modulate retrotransposition. *Nucleic Acids Res* 2013; 41: 575–85.
136. Dai L, Taylor MS, O'Donnell KA, Boeke JD. Poly(A) binding protein C1 is essential for efficient L1 retrotransposition and affects L1 RNP formation. *Mol Cell Biol* 2012; 32: 4323–36.
137. Stetson DB, Ko JS, Heidmann T, Medzhitov R. Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* 2008; 134: 587–98.
138. Carreira PE, Richardson SR, Faulkner GJ. L1 retrotransposons, cancer stem cells and oncogenesis. *FEBS J* 2014; 281: 63–73.
139. Goodier JL. Retrotransposition in tumors and brains. *Mob DNA* 2014; 5: 11.
140. Tubio JMC, Li Y, Ju YS, Martincorena I, Cooke SL, Tojo M, Gundem G, Pipinikas CP, Zamora J, Raine K, Menzies A, Roman-Garcia P, Fullam A, Gerstung M, Shlien A, Tarpey PS, Papaemmanuil E, Knappskog S, Van Loo P, Ramakrishna M, Davies HR, Marshall J, Wedge DC, Teague JW, Butler AP, Nik-Zainal S, Alexandrov L, Behjati S, Yates LR, Bolli N, Mudie L, Hardy C, Martin S, McLaren S, O'Meara S, Anderson E, Maddison M, Gamble S, Foster C, Warren AY, Whitaker H, Brewer D, Eeles R, Cooper C, Neal D, Lynch AG, Visakorpi T, Isaacs WB, van't Veer L, Caldas C, Desmedt C, Sotiriou C, Aparicio S, Foekens JA, Eyfjord JE, Lakhani SR, Thomas G, Myklebost O, Span PN, Borresen-Dale A-L, Richardson AL, Van de Vijver M, Vincent-Salomon A, Van den Eynden GG, Flanagan AM, Futreal PA, Janes SM, Bova GS, Stratton MR, McDermott U, Campbell PJ. Extensive transduction of non-repetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* 2014; 345: 1251343.
141. Consortium TEP. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; 489: 57–74.
142. Park PJ. [ChIP-Seq: advantages and challenges of a maturing technology.](#) *Nat Rev Genet* 2009; 10: 669–80.
143. Taslim C, Huang K, Huang T, Lin S. Analyzing ChIP-seq data: preprocessing, normalization, differential identification, and binding pattern characterization. *Methods Mol Biol* 2012; 802: 275–91.
144. Day DS, Luquette LJ, Park PJ, Kharchenko PV. [Estimating enrichment of repetitive elements from high-throughput sequence data.](#) *Genome Biol* 2010; 11: R69.
145. Thomas A Le, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, To KF. [Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state.](#) *Genes Dev* 2013; 27: 390–9.
146. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. [Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.](#) *Nat Protoc* 2012; 7: 562–78.
147. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011; 12: 323.
148. Chung D, Kuan PF, Li B, Sanalkumar R, Liang K, Bresnick EH, Dewey C, Keles S. Discovering transcription factor binding sites in highly repetitive regions of genomes with multi-read analysis of ChIP-Seq data. *PLoS Comput Biol* 2011; 7: e1002111.
149. Wang R, Hsu H-K, Blattler A, Wang Y, Lan X, Wang Y, Hsu P-Y, Leu Y-W, Huang TH-M, Farnham PJ, Jin VX. Locating non-unique matched tags (LONUT) to improve the detection of the enriched regions for ChIP-seq data. *PLoS One* 2013; 8: e67788.

150. Fang X, Nevo E, Han L, Levanon EY, Zhao J, Avivi A, Larkin D, Jiang X, Feranchuk S, Zhu Y, Fishman A, Feng Y, Sher N, Xiong Z, Hankeln T, Huang Z, Gorbunova V, Zhang L, Zhao W, Wildman DE, Xiong Y, Gudkov A, Zheng Q, Rechavi G, Liu S, Bazak L, Chen J, Knisbacher BA, Lu Y, Shams I, Gajda K, Farré M,

Kim J, Lewin HA, Ma J, Band M, Bicker A, Kranz A, Mattheus T, Schmidt H, Seluanov A, Azpurua J, McGowen MR, Ben Jacob E, Li K, Peng S, Zhu X, Liao X, Li S, Krogh A, Zhou X, Brodsky L, Wang J. Genome-wide adaptive complexes to underground stresses in blind mole rats *Spalax*. *Nat Commun* 2014; 5: 3966.



Maxime Bodak is a PhD student in biology at ETH Zurich, Molecular Life Sciences program. He holds a master degree in Genetics from the “Magistère Européen de Génétique” – University of Paris VII, Paris, France. In 2013, he joined Prof. Constance Ciaudo’s lab at the Institute of Molecular health Science, ETH Zurich and now works on LINE-1 regulation using the mouse embryonic stem cells as model.



Jian Yu is a PhD student in bioinformatics at ETH Zurich, Molecular and Translational Biomedicine program. He got his master degree in Tongji University, Shanghai, China. In 2013, he joined Prof. Constance Ciaudo’s lab at the Institute of Molecular health Science, ETH Zurich and now works on data analysis of next generation sequencing data.



Constance Ciaudo holds a PhD degree in Genetics from the University of Paris VII, Paris, France. After four years as a postdoctoral fellow at the Curie Institute in Paris, supervised by Prof. Edith Heard and at the Institute of Plant Biology (IBMP) in Strasbourg, France, supervised by Prof. Olivier Voinnet, she worked from 2010–2013 at ETH Zurich, Switzerland. She earned recognition for her Postdoctoral research in the field of RNA by receiving a FEBS Distinguished Young Investigator Award in 2013. She has been Assistant Professor of RNAi and Genome Integrity at the Institute of Molecular Health Science at ETH Zurich since April 2013.